The Latent-First Laboratory: A Manifesto for Efficient, Audit-Based AI Science

Anonymous Authors[1]

Abstract

High-throughput scientific workflows generate petabytes of visual data, historically mandating fully decompressed, human-readable formats. However, standard "decode-then-classify" downstream analysis introduces severe computational and storage bottlenecks. As scientific discovery increasingly relies on autonomous AI agents, we argue that maintaining entire datasets in pixel space is an unsustainable artifact of human-centric science. In this position paper, we propose the AI-Native Repository, a paradigm shift where the default state of scientific data is a compressed latent representation optimized for machine comprehension. Prior work suggests that learned or task-oriented compressed representations can preserve much of the semantic information needed for downstream inference, supporting the feasibility of latent-first pipelines. To address the trust gap between human researchers and autonomous systems, we also propose a Human Audit Slice policy, in which a small fraction of uncompressed data is preserved for manual verification and governance. Ultimately, we advocate for a structural shift toward machine-native data architectures, establishing new community standards for efficient, scalable, and verifiable AI-driven science.

1. Introduction

The era of big data in the biological and physical sciences has reached a critical inflection point. As demonstrated by recent advancements in high-speed volumetric imaging, such as PetaKit5D (Ruan et al.), modern light-sheet microscopy is now capable of generating data at rates exceeding four terabytes per hour, resulting in single experiments that scale into the petabyte range. While PetaKit5D and similar frameworks have made heroic strides in optimizing the "decode-then-process" pipeline—achieving order-of-magnitude improvements in Tiff/Zarr I/O, deskewing, and deconvolution—they remain tethered to a fundamental architectural vestige: the requirement that data be reconstructed into human-readable pixels before scientific analysis can begin (Ruan et al.).

We argue that this "pixel-first" requirement is the primary bottleneck hindering the next generation of autonomous discovery and that scientific computing should no longer assume pixel space as the default canonical representation. Instead, we propose a Latent-First Laboratory model in which raw observations are mapped directly into compressed, task-oriented latent representations. The key idea is simple: if the next stage of analysis is an AI system, then the stored representation should be optimized for machine comprehension rather than human readability. Recent work supports the feasibility of this shift. In climate and weather science, learned latent compression can preserve performance on downstream tasks while substantially reducing data volume (Han et al.). In high-throughput imaging, learned representations have improved downstream profiling performance and computational efficiency (Moshkov et al.). More broadly, latent-space imaging methods show that compressed representations can support downstream tasks directly, without requiring full reconstruction.

At the same time, a latent-first infrastructure introduces a legitimate trust problem. Scientists need to inspect, audit, and reproduce results, especially in high-stakes domains where model errors can propagate silently. For that reason, we propose a Human Audit Slice policy: a small fraction of the original uncompressed data should always be preserved, stratified across time, modality, batch, and uncertainty regime, so human researchers can verify model behavior and perform independent checks. This is consistent with broader AI governance thinking, which emphasizes layered oversight, documented access, and human judgment as the final arbiter.

This paper makes three important contributions:

1. It defines the latent-first repository as a machine-native scientific data architecture.

2. It introduces the Human Audit Slice as a governance mechanism for preserving human verifiability under compression.

[1]Anonymous Institution, Anonymous City, Anonymous Region, Anonymous Country. Correspondence to: Anonymous Author <anon.email@domain.com>.

Submitted to the AI for Science workshop (ICML 2026).

3. It proposes an evaluation agenda for determining when latent-first storage is scientifically valid, including rare-event recall, distribution-shift calibration, audit yield, and future-task robustness.

2. Related Work

Our proposal sits at the intersection of three existing research streams: petabyte-scale scientific imaging, learned representation and compression, and human-centered AI governance. While these areas have each advanced substantially, they have largely evolved under different assumptions about what scientific data should look like, who the primary consumer of that data is, and where the burden of interpretation should lie.

First, a growing body of work has addressed the practical bottlenecks of large-scale scientific imaging. In light-sheet microscopy and related volumetric modalities, modern acquisition systems now generate data at rates that quickly exceed the capacity of conventional storage and analysis pipelines (Ruan et al.). Systems such as PetaKit5D and related tools have focused on accelerating the "decode-then-process" workflow by improving deskewing, deconvolution, stitching, and I/O performance. These contributions are highly important, but they still assume that the canonical scientific object is a fully reconstructed pixel volume. In other words, they optimize the existing architecture rather than questioning whether pixel-space reconstruction should remain the default storage and analysis format. Our position paper differs by arguing that, once downstream consumers are autonomous machine systems, the repository itself should be redesigned around machine-native representations rather than human-readable pixels.

Second, our proposal is informed by research on learned compression, latent representations, and representation learning. Across computer vision and scientific machine learning, compressed embeddings have been shown to preserve substantial semantic information while dramatically reducing dimensionality. Recent work on latent compression in scientific contexts suggests that compact representations can support downstream prediction and inference without requiring full reconstruction of the original signal (Liu et al.; Han et al.). Related work in image-based profiling and representation learning further demonstrates that biologically meaningful structure can be captured in learned feature spaces (Moshkov et al.). These results support the technical plausibility of latent-first pipelines, but most prior work treats latent spaces as intermediate model artifacts rather than as the canonical stored form of scientific data. We take the stronger view that latent representations should become first-class objects in scientific repositories, with reconstruction reserved for selective audit and visualization.

A related thread comes from work on compression, denoising, and scalable image processing. In imaging and signal processing, latent-space methods have been used not only to reduce storage but also to improve downstream performance by discarding noise and preserving task-relevant structure (Alvar et al.; Testolina et al.). Similarly, recent latent-space imaging work suggests that compressed representations can support direct downstream use in settings where full fidelity reconstruction is unnecessary (Souza et al.). These approaches reinforce the idea that compression need not be a purely archival operation. However, our framework extends this logic from isolated model pipelines to the level of data infrastructure: instead of compressing individual inputs for a specific model, we propose repositories whose default state is already latent and searchable.

Third, our Human Audit Slice proposal is informed by work on AI governance, auditing, and human oversight. As AI systems become more capable and more embedded in scientific and operational settings, concerns about transparency, accountability, and silent failure have become central. Prior work on AI auditing and digital accountability argues that machine-driven systems require explicit oversight structures rather than blind trust (Birhane et al.). Broader governance discussions likewise emphasize the importance of human intervention, auditability, and procedural checks in high-stakes AI workflows (Art). Our Human Audit Slice extends this perspective to scientific data infrastructure by preserving a small but strategically selected subset of raw data for manual review. Rather than requiring all data to remain human-readable, we argue that accountability can be achieved through selective preservation, uncertainty-aware sampling, and governance policies that retain access to the original signal when needed.

Finally, our proposal is motivated by the growing use of autonomous and agentic AI systems in scientific discovery. Recent progress in scientific foundation models, generative modeling, and AI-assisted discovery suggests that machine systems are increasingly capable of acting as primary interpreters of complex experimental data (Gridach et al.). As this shift continues, the dominant assumption that scientific data must first be made visually interpretable for humans becomes less compelling. The AI-Native Repository is therefore best understood not as a compression trick, but as an architectural response to a changing scientific workflow: one in which machines, not humans, are the immediate consumers of most data, while humans retain oversight through structured audit mechanisms.

Taken together, prior work establishes the feasibility of both latent representations and human oversight, but it does not yet unify them into a coherent repository design. Our contribution is to connect these threads into a single position: scientific infrastructure should move from pixel-first storage toward latent-first repositories, while preserving a principled Human Audit Slice for verification, governance, and reproducibility.

3. The Case for Latent-First Science

The central limitation of pixel-first science is not only storage size, but the repeated cost of reconstructing information that machine models often do not need. In PetaKit5D, even optimized image handling still has to contend with expensive I/O and transformation steps; the benchmarked Tiff read/write paths in MATLAB take on the order of minutes for large volumes, and the paper reports that these slow I/O speeds are a “considerable bottleneck” for efficient processing. The broader implication is that “fast” deconvolution and deskewing remain expensive when the acquisition scale is measured in terabytes per hour and experiments reach petabyte scale.

This overhead matters because raw pixels contain much more than biological signal. High-frequency noise, sensor artifacts, and redundant background structure may be useful for calibration, but they are often unnecessary for downstream tasks such as segmentation, phenotype detection, or localization-based classification. A learned latent representation can act as a semantic filter: instead of storing every pixel equally, it preserves the task-relevant degrees of freedom needed for machine inference. This is the same intuition behind recent work on latent compression and representation learning, where compact encodings can retain enough structure for downstream prediction and analysis.

There is also a sustainability argument. Scientific computing is increasingly entangled with energy-intensive infrastructure, and data centers are already a significant and growing electricity burden. If every experiment must be fully decompressed, deskewed, deconvolved, and written back out as a conventional image stack before analysis can begin, the system pays a large recurring cost in compute, cooling, and storage. Latent-first science offers a practical “Green AI” direction: reduce the volume of data that must be kept in full pixel form, reduce the amount of repeated decoding, and reduce the carbon footprint of large-scale experimental pipelines.

Finally, latent-first inference is essential for autonomous loops. If an AI agent is controlling a microscope in real time, it cannot afford to wait for a long file-write and reconstruction cycle before deciding whether to refocus, re-scan, or terminate acquisition. PetaKit5D’s reported multi-minute Tiff operations make clear that conventional image pipelines are too slow for closed-loop autonomy at scale. A direct latent pipeline can support much faster feedback, enabling millisecond- to second-level decisions instead of minute-level delays.

4. Proposed Framework: The AI-Native Repository

We propose the AI-Native Repository as a new default architecture for large-scale scientific data. In this model, the acquisition system feeds data into a Latent Encoder that transforms the signal into a compact representation immediately, before the data is written to long-term storage. Rather than treating compression as an afterthought, the repository treats the latent code as the canonical stored object, with pixel-space reconstruction available only when needed for audit, visualization, or specialized downstream use.

This repository functions as a searchable vector database for scientific observations. An autonomous agent can query the repository semantically—for example: “find all cells with mitotic defects,” or “return volumes likely to contain segmentation failures”—without decompressing the full dataset. This shifts the repository from a passive storage volume to a machine-native environment optimized for retrieval, similarity search, and uncertainty estimation.

4.1. The Adaptive Audit Policy

To ensure scientific integrity and accountability, we propose a Human Audit Slice (HAS) protocol. Rather than a static retention rate, the HAS is an adaptive audit policy where the fraction of raw data preserved is a function of task risk, phenotype rarity, model uncertainty, distribution shift, and downstream consequences.

The policy is defined by the following logic:

$$
\begin{aligned}
HAS = \text{Baseline} &+ \text{Uncertainty Trigger} \\
&+ \text{Rare-Event Escalation} \\
&+ \text{Risk-Based Zones}
\end{aligned} \tag{1}
$$

This multi-tiered approach ensures that while the majority of routine data remains in latent form, human verification is focused where it provides the highest marginal value. This makes the framework resilient to black swan events or model failures that a fixed sampling rate might miss.

4.2. Risk-Based Retention Tiers

The exact fraction of raw data retained is adjustable based on the specific scientific or clinical context. Table 1 outlines the suggested retention policy across different experimental workflows. It is important to note that these values represent starting heuristics for community discussion; we anticipate that rigorous ablation studies on rare-event detection will be required to formalize these thresholds as scientific standards.

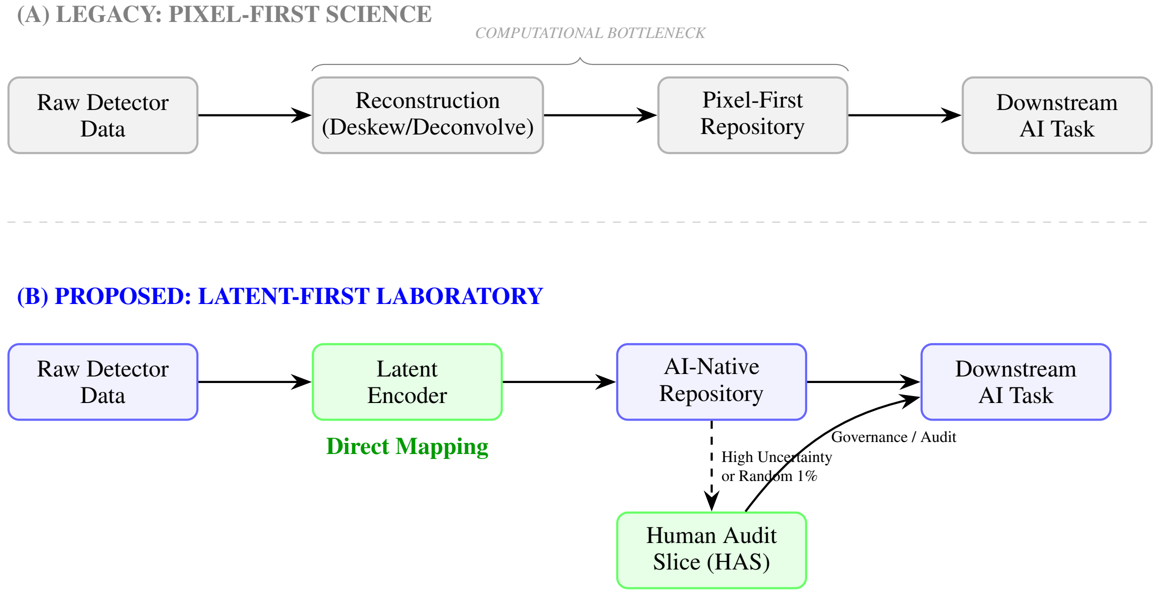

Figure 1. **Architectural shift from Pixel-First to Latent-First science.** (A) Conventional workflows mandate expensive pixel-space reconstruction and storage, creating a severe bottleneck for petabyte-scale experiments. (B) Our proposed AI-Native Repository maps raw data directly to latents, bypassing traditional reconstruction for the vast majority of data while preserving a selective **Human Audit Slice (HAS)** for verifiability and governance.

Table 1. Suggested raw data retention tiers based on workflow risk.

Workflow Risk	**Suggested Retention**	**Rationale**
Routine Screening	1–5%	Low-risk, high-volume calibration.
Rare-Event Discovery	10–50% (stratified)	Targeted retention of novel phenotypes.
Clinical/Safety-Critical	∼100%	Regulatory and diagnostic audit.
Out-of-Distribution	Full Retention	Immediate review until recalibration.

5. Technical Challenges and Open Questions

The first challenge is forward compatibility. Latents that are optimal for today's models may become suboptimal as architectures, tasks, and scientific questions evolve. A repository designed around a 2026 model might not be ideal for a 2036 model. This suggests that latent-first science must include versioning, model provenance, and periodic re-encoding strategies so that archived data does not become locked into a historical representation. In other words, latent storage should be treated as dynamic infrastructure, not a one-time compression event.

The second challenge is standardization. If every lab develops its own latent space, interoperability will collapse. We therefore need community standards for scientific latent representations, analogous to how RGB, TIFF, and Zarr standardized parts of image exchange. A shared latent protocol would need agreed metadata, calibration rules, uncertainty flags, and decoding interfaces. Without such standards, latent-first repositories may become isolated silos rather than shared scientific infrastructure.

The third challenge is the black box risk. A latent representation can be efficient precisely because it discards information, but that creates the possibility of silent information loss. A rare phenotype, subtle artifact, or previously unknown biological feature could be compressed away before anyone realizes it matters. This is why the Human Audit Slice matters: it provides a direct check against overcompression and model myopia. It also suggests that latent-first systems should be evaluated not only on task accuracy, but also on recovery of rare events, calibration under distribution shift, and reproducibility across labs.

6. A Workflow Example

To make the AI-Native Repository concrete, consider a representative high-throughput light-sheet microscopy experiment. A specimen is imaged continuously over time, producing a large volumetric stream of raw measurements. In a conventional pipeline, these measurements are first written to disk, then reconstructed into pixel-space volumes through deskewing, deconvolution, stitching, and related preprocessing steps, and only then passed to downstream analysis tools. Even when these preprocessing tools are optimized, the workflow still requires the system to materialize a full human-readable image stack before machine analysis can begin (Ruan et al.). Under our proposed framework, this sequence is inverted.

In the Latent-First Laboratory, the microscope output is immediately passed through a Latent Encoder that transforms each incoming frame or volume into a compact representation. The latent code is stored as the primary archival object in the repository, along with metadata describing acquisition conditions, encoder version, model confidence, and task-specific annotations. Rather than treating compression as a preprocessing step, the repository treats the latent code as the canonical scientific record. This design reduces storage overhead while preserving the information needed for machine inference and retrieval (Liu et al.; Han et al.; Souza et al.).

Now consider a downstream task such as mitotic defect detection. An autonomous agent queries the repository for latent embeddings corresponding to cells whose morphology or motion deviates from expected patterns. Because the repository is searchable in latent space, the agent does not need to reconstruct every volume in full pixel form. Instead, it can retrieve only those samples whose latent structure matches the target phenotype or whose uncertainty score indicates a likely anomaly. If the model is highly confident, the latent result may be sufficient to update the agent's state and continue acquisition. If the model is uncertain, the corresponding sample is routed into the Human Audit Slice for review.

The Human Audit Slice is the mechanism that preserves scientific accountability. Suppose the latent encoder detects a rare or ambiguous event: a division abnormality, an unusual protein localization pattern, or a possible imaging artifact. Rather than trusting the latent prediction blindly, the system preserves the raw pixel-space version of that specific frame or subvolume for human inspection. The sample is selected using uncertainty-aware criteria, not arbitrary random subsampling. This allows the repository to focus manual review on the cases most likely to reveal model failure, distribution shift, or biologically interesting edge conditions. In effect, the audit slice acts as a safety valve that balances efficiency with verifiability (Birhane et al.; Gu et al.).

A practical implementation might preserve only a small fraction of the data in raw form, such as one percent of frames or volumes, while retaining the remaining ninety-nine percent only as latent codes. The preserved slice could be stratified by experimental condition, time point, batch, and uncertainty score to ensure that it is scientifically representative rather than merely convenient. This subset would support spot checks, reproducibility audits, and human-led discovery without forcing the entire dataset to remain in a costly pixel-first format.

This workflow also enables closed-loop acquisition. If the repository and agent detect that a region of interest is producing high-value observations, the microscope can be instructed to focus on that region, increase temporal resolution, or trigger higher-fidelity capture. Because decisions are made directly from latent states rather than from fully reconstructed image stacks, the feedback loop can operate much faster than a traditional decode-then-classify pipeline. The result is not only lower storage cost, but also a more responsive scientific instrument that can adapt acquisition in real time.

This example illustrates the broader logic of the AI-Native Repository. The goal is not to eliminate pixel-space imaging altogether, but to stop treating it as the default canonical representation of scientific data. By making latent state the primary stored object and pixel-space reconstruction an exception reserved for auditing and special use cases, the repository aligns the data architecture with the needs of autonomous scientific systems.

7. Conclusion

Scientific infrastructure should stop treating AI as a human with faster eyes. Autonomous systems do not need every dataset to be laid out in pixel space before they can reason about it, and insisting on that assumption imposes a large and increasingly unnecessary tax on science. The next phase of discovery requires data architectures that are native to machine inference from the start.

The Latent-First Laboratory is our proposed answer: encode observations into compact representations, run discovery in latent space, and preserve a governed Human Audit Slice for accountability. If science is to scale into the exabyte era, it will need repositories built for machines first and humans second, rather than the reverse.

Impact Statement

This work proposes a shift toward AI-native scientific data architectures. The potential benefits are significant: by reducing data volume by orders of magnitude, we lower the energy footprint of big data science (addressing the environmental impact of data centers) and democratize high-throughput research for labs without petabyte-scale storage budgets.

However, we acknowledge the risk of compression models inadvertently discard subtle, novel biological signals that do not fit the training distribution. This could lead to a silent discovery gap where rare phenomena are never recorded. To mitigate this, our proposed Human Audit Slice (HAS) serves as a governance initiative to ensure human-in-the-loop accountability.

Crucially, the retention percentages suggested in Table 1 are preliminary heuristics based on current imaging bottlenecks. A primary uncertainty remains the minimum viable retention needed to maintain scientific reproducibility across different modalities. We advocate for future research to empirically stress-test these thresholds using rare-event recall benchmarks.

References

Article 14: Human oversight | EU artificial intelligence act. URL https://artificialintelligenceact.eu/article/14/.

Alvar, S. R., Ulhaq, M., Choi, H., and Bajić, I. V. Joint image compression and denoising via latent-space scalability. URL http://arxiv.org/abs/2205.01874.

Birhane, A., Steed, R., Ojewale, V., Vecchione, B., and Raji, I. D. AI auditing: The broken bus on the road to AI accountability. URL http://arxiv.org/abs/2401.14462. version: 1.

Gridach, M., Nanavati, J., Abidine, K. Z. E., Mendes, L., and Mack, C. Agentic AI for scientific discovery: A survey of progress, challenges, and future directions. URL http://arxiv.org/abs/2503.08979.

Gu, H., Schreyer, M., Moffitt, K., and Vasarhelyi, M. Artificial intelligence co-piloted auditing. 54:100698. ISSN 1467-0895. doi: 10.1016/j.accinf.2024.100698. URL https://www.sciencedirect.com/science/article/pii/S1467089524000319.

Han, T., Chen, Z., Guo, S., Xu, W., Ouyang, W., and Bai, L. Climate science data can be compressed efficiently by dual-stage extreme compression with a variational auto-encoder transformer. 6 (1):955. ISSN 2662-4435. doi: 10.1038/s43247-025-02903-z. URL https://www.nature.com/articles/s43247-025-02903-z.

Liu, Q., Gong, B., Zhuang, X., Zhong, X., Kang, Z., and Li, H. Compressing high-resolution data through latent representation encoding for downscaling large-scale AI weather forecast model. URL http://arxiv.org/abs/2410.09109.

Moshkov, N., Bornholdt, M., Benoit, S., Smith, M., McQuin, C., Goodman, A., Senft, R. A., Han, Y., Babadi, M., Horvath, P., Cimini, B. A., Carpenter, A. E., Singh, S., and Caicedo, J. C. Learning representations for image-based profiling of perturbations. 15(1):1594. ISSN 2041-1723. doi: 10.1038/s41467-024-45999-1. URL https://www.nature.com/articles/s41467-024-45999-1.

Ruan, X., Mueller, M., Liu, G., Görlitz, F., Fu, T.-M., Milkie, D. E., Lillvis, J. L., Kuhn, A., Gan Chong, J., Hong, J. L., Herr, C. Y. A., Hercule, W., Nienhaus, M., Killilea, A. N., Betzig, E., and Upadhyayula, S. Image processing tools for petabyte-scale light sheet microscopy data. 21(12):2342–2352. ISSN 1548-7105. doi: 10.1038/s41592-024-02475-4. URL https://www.nature.com/articles/s41592-024-02475-4.

Souza, M., Zheng, Y., Kang, K., Mishra, Y. N., Fu, Q., and Heidrich, W. Latent space imaging. In *2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 28295–28305. doi: 10.1109/CVPR52734.2025.02635. URL https://ieeexplore.ieee.org/document/11094851. ISSN: 2575-7075.

Testolina, M., Upenik, E., and Ebrahimi, T. Towards image denoising in the latent space of learning-based compression. In *Proceedings of SPIE*. SPIE. doi: 10.1117/12.2597828. URL https://infoscience.epfl.ch/handle/20.500.14299/181174.